

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 17:14:36 ; Search time 74.2239 Seconds
(without alignments)
504.414 Million cell updates/sec

Title: US-09-807-933B-11

Perfect score: 1895

Sequence: 1 MKFSIIIASALLLAASSTYAA.....TFKAVTCAPIIAKTGCERK 346

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA.*

- 1: /cgn2_6/protdata/2/pubpaa/US08_NEW PUB.pdb.*
- 2: /cgn2_6/protdata/2/pubpaa/PCT_NEW PUB.pdb.*
- 3: /cgn2_6/protdata/2/pubpaa/US06_NEW PUB.pdb.*
- 4: /cgn2_6/protdata/2/pubpaa/US06_PUBCOMB.pdb.*
- 5: /cgn2_6/protdata/2/pubpaa/US07_NEW PUB.pdb.*
- 6: /cgn2_6/protdata/2/pubpaa/US07_PUBCOMB.pdb.*
- 7: /cgn2_6/protdata/2/pubpaa/US07_PUBCOMB.pdb.*
- 8: /cgn2_6/protdata/2/pubpaa/US08_PUBCOMB.pdb.*
- 9: /cgn2_6/protdata/2/pubpaa/US09_NEW PUB.pdb.*
- 10: /cgn2_6/protdata/2/pubpaa/US09_PUBCOMB.pdb.*
- 11: /cgn2_6/protdata/2/pubpaa/US10_NEW PUB.pdb.*
- 12: /cgn2_6/protdata/2/pubpaa/US10_PUBCOMB.pdb.*
- 13: /cgn2_6/protdata/2/pubpaa/US60_NEW PUB.pdb.*
- 14: /cgn2_6/protdata/2/pubpaa/US60_PUBCOMB.pdb.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	696.5	36.8	201	9	US-09-261-329-4
2	686.5	36.2	299	9	US-10-007-521-12
3	683.5	36.1	203	9	US-09-261-329-9
4	683.5	36.1	222	9	US-10-007-521-14
5	683.5	36.1	294	9	US-10-007-521-24
6	680	35.9	349	9	US-10-007-521-10
7	679.5	35.9	201	9	US-09-261-329-5
8	678.5	35.8	225	9	US-10-007-521-2
9	678.5	35.8	297	9	US-10-007-521-4
10	678.5	35.8	308	9	US-10-007-521-6
11	672.5	35.5	205	9	US-09-261-329-7
12	668	35.3	310	9	US-10-007-521-22
13	664	34.5	202	9	US-09-261-329-3
14	652	34.4	305	10	US-09-735-787-2
15	651.5	34.4	235	1	US-08-841-636A-31
16	651	34.0	202	9	US-09-261-329-1
17	644.5	34.0	376	10	US-09-735-787-4
18	640.5	33.8	203	9	US-09-261-329-6
19	634.5	33.5	226	9	US-10-007-521-16

20	634.5	33.5	293	9	US-10-007-521-20	Sequence 20, Appl
21	634.5	33.5	298	9	US-10-007-521-18	Sequence 18, Appl
22	629.5	33.2	203	9	US-09-261-329-8	Sequence 8, Appl
23	620	32.7	295	9	US-10-007-521-8	Sequence 8, Appl
24	616	32.5	202	9	US-09-261-329-2	Sequence 2, Appl
25	468.5	24.7	235	9	US-09-261-329-10	Sequence 10, Appl
26	456.5	24.1	211	9	US-09-261-329-11	Sequence 11, Appl
27	388	20.5	138	9	US-10-007-521-26	Sequence 26, Appl
28	230	12.1	2750	9	US-10-123-155-85	Sequence 85, Appl
29	228.5	12.1	18636	9	US-10-073-913-17	Sequence 17, Appl
30	227	12.0	2316	9	US-10-123-155-69	Sequence 69, Appl
31	227	12.0	3060	9	US-10-184-644-337	Sequence 337, Appl
32	227	12.0	3060	9	US-10-184-634-337	Sequence 337, Appl
33	226.5	12.0	493	9	US-10-197-294A-2	Sequence 2, Appl
34	223	11.8	4060	9	US-10-123-155-197	Sequence 197, Appl
35	222.5	11.7	4440	9	US-10-174-590-525	Sequence 525, Appl
36	222.5	11.7	4440	9	US-10-176-758-525	Sequence 525, Appl
37	222.5	11.7	4440	9	US-10-175-737-525	Sequence 525, Appl
38	222.5	11.7	4440	9	US-10-173-706-525	Sequence 525, Appl
39	222.5	11.7	4440	9	US-10-175-738-525	Sequence 525, Appl
40	222.5	11.7	4440	9	US-10-175-752-525	Sequence 525, Appl
41	222.5	11.7	4440	9	US-10-176-482-525	Sequence 525, Appl
42	222.5	11.7	4440	9	US-10-176-757-525	Sequence 525, Appl
43	222.5	11.7	4440	9	US-10-176-913-525	Sequence 525, Appl
44	222.5	11.7	4440	9	US-10-180-552-525	Sequence 525, Appl
45	222.5	11.7	4440	9	US-10-180-557-525	Sequence 525, Appl

ALIGNMENTS

RESULT 1

US-09-261-329-4
; Sequence 4, Application US/09261329
; Publication No. US20030092097A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Kim
; APPLICANT: Schuelein, Martin
; APPLICANT: Christiansen, Lars
; APPLICANT: Damgaard, Bo
; APPLICANT: Von Der Osten, Claus
; TITLE OF INVENTION: Cellulase Variants
; FILE REFERENCE: 4887.204-US
; CURRENT APPLICATION NUMBER: US/09/261,329
; CURRENT FILING DATE: 1999-03-03
; EARLIER APPLICATION NUMBER: 1013/96
; EARLIER FILING DATE: 1996-09-17
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Cellulase variants
US-09-261-329-4

Query Match 36.8%; Score 696.5; DB 9; Length 201;

Best Local Similarity 59.6%; Pred. No. 1.5e-47;

Matches 121; Conservative 28; Mismatches 51; Indels 3; Gaps 1;

QY	143	GNGRTTRYWDCCKPCSCAMDGKASVTKPVLTCAKDGVSRIGSDVQSCGCGQAYMCDNOP	202
DB	1	GSCKSTRYWDCKPCSCAMSGKASVNRPLVACDANNPLNDANVKSCKDCGGSAYTCANNSP	60
QY	203	WVNNDLVAFRAASISGASAFCCGCVLFTNTAVAGKFFVQVTTNGDLSNTHFD	262
DB	61	WAVNNLAFRAATKLSGGTSSWCACALYFTSGPVSCKTLVQVSTSGDLSNTHFD	120
QY	263	LQMPGGGVYFNGCOSQMTNTDGMGARYGGISSTSECCKLPTQOAGCKWRFQFNAD	322
DB	121	LNPFGGVSLFGCKREFQGLP---GAQYGGISRSSECDSPFAALKPGCQWRFQFNAD	177
QY	323	NPEVTFKAVTCAPIIAKTGCER	345

Wed Jun 18 17:55:03 2003

Db 178. NPEFTFKVQVCPSELTSRTGCKR 200

RESULT 2

US-10-007-521-12

Sequence 12, Application US/10007521

Publication No. US20030054539A1

GENERAL INFORMATION:

APPLICANT: Schulein, Martin

Andersen, Lene N.

Lassen, Soren F.

Kauppinen, Markus S.

Lange, Lene

Nielsen, Ruby I.

Ihara, Michiko

Takegi, Shinobu

TITLE OF INVENTION: No. US20030054539A1el Endoglucanases

NUMBER OF SEQUENCES: 109

CORRESPONDENCE ADDRESSES:

ADDRESSEE: No. US20030054539A1o No. US20030054539A1disk of No. US20030054539A1

STREET: 405 Lexington Avenue, 64th Floor

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/007,521

FILING DATE: 10-Dec-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/651,136

FILING DATE: 21-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J.

REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 4366.200-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 299 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLSCULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-10-007-521-12

Query Match 36.2%; Score 686.5; DB 9; Length 299;

Best Local Similarity 53.6%; Pred. No. 1.4e-46;

Matches 125; Conservative 31; Mismatches 66; Indels 11; Gaps 2;

QY 113 KTTTPTTKTTAAATTTSSSTGYSPIGSGFNGRTRTYWDCCKPSCAWDGKASVTKPVLT 172

Db 2 RSTPVLRTTLAALPL-----VASAASGSGSTRTYWDCCKPSCAWPGKAASQPVYA 53

QY 173 CAKDGVSRLGSDVQSGCGVQAYMNDNOPVWVNDLDTLQAGCKWRFKFNAD 232

Db 54 CDANFQRLSDFNVSQSGNGSAYSCADQTPFVAVNDNLAYGFAATSIAGSSSSWCCACYA 113

QY 233 LTFNTAVAGKFFVQVVTNTGDDLSNTHFDLQMPGGVGYFNGCQSQMNTNTDGMARYG 292

Db 114 LTFISGVAGKTNVWVQSTGGLDLSNTHFDLQMPGGVGYFNGCQSQMNTNTDGMARYG 170

QY 293 GISSISBCDKLPTQLQAGCKWRFKFNADNPEVTKAVTTPCAEIIKTGCR 345

Db 171 GISSRQDCDSFPAPLPGCQWRDQFONADNPTFTFQQVQCPAEIIVARSGCKR 223

Db 178. NPEFTFKVQVCPSELTSRTGCKR 200

RESULT 3

US-09-261-329-9

Sequence 9, Application US/09261329

Publication No. US20030092097A1

GENERAL INFORMATION:

APPLICANT: Andersen, Kim

APPLICANT: Schulein, Martin

APPLICANT: Christiansen, Lars

APPLICANT: Damgaard, Bo

APPLICANT: Von Der Osten, Claus

TITLE OF INVENTION: Cellulase Variants

FILE REFERENCE: 4887.204-US

CURRENT APPLICATION NUMBER: US/09/261,329

EARLIER FILING DATE: 1999-03-03

EARLIER APPLICATION NUMBER: 1013/96

NUMBER OF SEQ ID NOS: 26

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 9

LENGTH: 203

TYPE: PRT

ORGANISM: Cellulase variants

US-09-261-329-9

Query Match 36.1%; Score 683.5; DB 9; Length 203;

Best Local Similarity 58.6%; Pred. No. 1.6e-46;

Matches 119; Conservative 30; Mismatches 51; Indels 3; Gaps 2;

QY 144 NGRTRYWDCCKPSCAWDGKASVTKPVLTCAKDGVSRLGSD-VQSGCGVQAYMNDNOP 202

Db 2 SGVTRYWDCCKPSCAWTGKASVKEVGTCDINDNAQTPTSDLLKSSDCGSAAYCSNQGP 61

QY 203 WVVNDLDTLQAGCKWRFKFNADNPEVTKAVTTPCAEIIKTGCR 345

Db 62 WVNDSLSYGFAAAKLSGKQETDWCQCYKLTFTSTAVSGKQMVQITNTGDLGNHFD 121

QY 263 LQMPGGVGYFNGCQSQMNTNTDGMARYGGISSISECDKLPQLQAGCKWRFKFNAD 322

Db 122 IAMPGGVGYFNGCQSQMNTNTDGMARYGGISSISECDKLPQLQAGCKWRFKFNAD 179

QY 323 NPEVTKAVTTPCAEIIKTGCR 345

Db 180 NPTVDWEPTVCPQELVARTGCSR 202

RESULT 4

US-10-007-521-14

Sequence 14, Application US/10007521

Publication No. US20030054539A1

GENERAL INFORMATION:

APPLICANT: Schulein, Martin

Andersen, Lene N.

Lassen, Soren F.

Kauppinen, Markus S.

Lange, Lene

Nielsen, Ruby I.

Ihara, Michiko

Takegi, Shinobu

TITLE OF INVENTION: No. US20030054539A1el Endoglucanases

NUMBER OF SEQUENCES: 109

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. US20030054539A1o No. US20030054539A1disk of No. US20030054539A1

STREET: 405 Lexington Avenue, 64th Floor

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/007,521
FILING DATE: 10-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/651,136
FILING DATE: 21-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4366,200
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 222 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-10-007-521-14

```

Query Match	36.1%;	Score 683.5;	DB 9;	Length 222;
Best Local Similarity	58.6%;	Pred. No. 1.8e-46;		
Matches 119;	Conservative 30;	Mismatches 51;	Indels 3;	Gaps 2;
Qy	144	NGRTTRWDCCKPSCAWDGRKASVTKPVLTKAKDGVSRGLGSD-VOSGCVGGQAYMGNNDNQP	202	
Db	21	SGVTRTRWDCCKPSCAWTGTGKASVSKPVGTCDINNDAGTSPDLLKSSCDGGSAYVCSNQGP	80	
Qy	203	WVYNDLDIAYGFAAASLGSAGASAPCCGCBELTFTNTAVAGKKFVVQVNTTGGDLSTNHFD	262	
Db	81	WAVNDSLSYGFAAAKLSGKOETDCCGCKYLFTTSTAVSGKQMTVQITNTGGDIGNHFD	140	
Qy	263	LQWPGGGVGVFNGCQSQWNTNTDQWAGARYGGISISSECDKLPTOLQAGCKWRFCQFKNAD	322	
Db	141	IAMPGGGVGVIFNGCSKQW--NGINLGNQYGGFTDRSQCATLPSKWKQASCNWRFDWFENAD	198	
Qy	323	NPEVTFKAVTCAPABIIAKTGCER	345	
Db	199	NPTVDWEPVTCPOELVARTGCSR	221	

```

RESULT 5
US-10-007-521-24
; Sequence 24, Application US/10007521
; Publication No. US20030054539A1
; GENERAL INFORMATION:
; APPLICANT: Schulein, Martin
; Andersen, Lene N.
; Lassen, Soren P.
; Kauppinen, Markus S.
; Lange, Lene
; Nielsen, Ruby I.
; Ihara, Michiko
; Takagi, Shinobu
; TITLE OF INVENTION: No. US20030054539A1el Endoglucanases
; NUMBER OF SEQUENCES: 109
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. US20030054539A1o No. US20030054539A1
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/007-521

```

```

? FILING DATE: 10-Dec-2001
? CLASSIFICATION: <Unknown>
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US/08/651,136
? FILING DATE: 21-MAY-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Lambiris, Elias J.
? REGISTRATION NUMBER: 33,728
? REFERENCE/DOCKET NUMBER: 4366.200-US
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 212-867-0123
? TELEFAX: 212-878-9655
? INFORMATION FOR SEQ ID NO: 24:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 294 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-10-007-521-24

Query Match          36.1%; Score 683.5; DB 9; Length 294;
Best Local Similarity 58.6%; Pred No. 2.4e-46;
Matches 119; Conservative 30; Mismatches 51; Indels 3; Gaps 2

Qy      144 NGRTRTYDWCCKPSCAWDGKASVTKPVLTCAKGVSRLGSD-VOSGCYGGQAYMCNDNQP 202
Db      :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy      203 WVNDDILAYGAFAAASLGSAGASAPCCCGCYELTFTNTAVAGKKFVVQVTNTGDLSLNHPD 262
Db      :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      81 WAVNDSLISYGFAAAKLSGKOETDWCCGCGYKLTFTSTAVSGKMIVQIITNTGGDLGNHFD 140
Qy      263 LQMPGGGVGFNGCQSOWNTNTDCMGARYGGISSISECDKLPLOAQACKRFGWFKXAD 322
Db      |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      141 IAMPGGVGIFNGCSKW--NGINLGNYGGFTDRSQATLPSKMQASCNRWDFWFENAD 198
Qy      323 NPEVTFKAVTCPAEIIARTGCR 345
Db      ||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      199 NPTVDMEFVTCQELVARTGCSR 221

RESULT 6
US-10-007-521-10
Sequence 10, Application US/10007521
Publication No. US20030054539A1
GENERAL INFORMATION:
APPLICANT: Schuelein, Martin
Andersen, Lene N.
Lassen, Soren P.
Kauppinen, Markus S.
Lange, Lene
Nielsen, Ruby I.
Ihara, Michiko
Takagi, Shinobu
TITLE OF INVENTION: No. US20030054539A1el Endoglycanases
NUMBER OF SEQUENCES: 109
CORRESPONDENCE ADDRESSES:
ADDRESSEE: No. US20030054539Alo No. US20030054539Adisk of No. US
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/007,521
FILING DATE: 10-Dec-2001
CLASSIFICATION: <Unknown>

```


Db	123	KKVIQATNTGGDLDGNHFDLAI	PGGVGIFNACTDQYGAPPNGWGDRYGIHSKECES	182
Qy	303	LPTQLQACKRRFGFKVADNP	EVTFKAVTCPAEIIAKTGCR	345
Db	183	FPEALKPCGNRWFQFNADNP	SVTFQEVACPSELTSKSGCSR	225

RESULT 10
US-10-007-521-6
; Sequence 6, Application US/10007521
; Publication No. US20030054539A1
; GENERAL INFORMATION:
; APPLICANT: Schulein, Martin
; ; Andersen, Lene N.
; ; Lassen, Soren F.
; ; Kauppinen, Markus S.
; ; Lange, Lene
; ; Nielsen, Ruby I.
; ; Ihara, Michiko
; ; Takagi, Shinobu
; TITLE OF INVENTION: No. US20030054539A1el Endoglucanases
; ;
; NUMBER OF SEQUENCES: 109
; ;
; CORRESPONDENCE ADDRESS:
; ;
; ADDRESSEE: No. US20030054539A1o No. US20030054539A1
; ;
; STREET: 405 Lexington Avenue, 64th Floor
; ;
; CITY: New York
; ;
; STATE: New York
; ;
; COUNTRY: United States of America
; ;
; ZIP: 10174-6401
; ;

```

1  ZIP: 10174-6401
2
3  COMPUTER READABLE FORM:
4
5  MEDIUM TYPE: Floppy disk
6
7  COMPUTER: IBM PC compatible
8
9  OPERATING SYSTEM: PC-DOS/MS-DOS
10
11  SOFTWARE: PatentIn Release #1.0, Version #1.30
12
13  CURRENT APPLICATION DATA:
14
15  APPLICATION NUMBER: US/10/007,521
16
17  FILING DATE: 10-Dec-2001
18
19  CLASSIFICATION: <Unknown>
20
21  PRIOR APPLICATION DATA:
22
23  APPLICATION NUMBER: US/08/651,136
24
25  FILING DATE: 21-MAY-1996
26
27  ATTORNEY/AGENT INFORMATION:
28
29  NAME: Lambiris, Elias J.
30
31  REGISTRATION NUMBER: 33,728
32
33  REFERENCE/DOCKET NUMBER: 4366.200-US
34
35  TELECOMMUNICATION INFORMATION:
36
37  TELEPHONE: 212-867-0123
38
39  TELEFAX: 212-878-9655
40
41  INFORMATION FOR SEQ ID NO: 6:
42
43  SEQUENCE CHARACTERISTICS:
44
45  LENGTH: 308 amino acids
46
47  TYPE: amino acid
48
49  TOPOLOGY: linear
50
51  MOLECULE TYPE: protein
52
53  SEQUENCE DESCRIPTION: SEQ ID NO: 6:
54
55  US-10-007-521-6

```

[illegible]

QY 303 LPTQLQAGCKWRFGWFKADNPVETFKAVTCTPAEIIAKTCER 345
Db 183 FPEALPGCNWRFDFWQADNPVETFKAVTCTPAEIIAKTCER 225

RESULT 11

US-09-261-329-7
; Sequence 7, Application US/09261329
; Publication No. US20030092097A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Kim
; APPLICANT: Schulein, Martin
; APPLICANT: Christiansen, Lars
; APPLICANT: Damgaard, Bo
; APPLICANT: Von Der Osten, Claus
; TITLE OF INVENTION: Cellulase Variants
; FILE REFERENCE: 4887.204-US
; CURRENT APPLICATION NUMBER: US/09/261,329
; CURRENT FILING DATE: 1999-03-03
; EARLIER APPLICATION NUMBER: 1013/96
; EARLIER FILING DATE: 1996-09-17
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Cellulase variants
US-09-261-329-7

Query Match 35.5%; Score 672.5; DB 9; Length 205;
Best Local Similarity 57.1%; Pred. No. 1.2e-45;
Matches 117; Conservative 23; Mismatches 62; Indels 3; Gaps 3;

QY 143 GNCETTRYWDCCKPSCAWDGKASVTKPVLTCAK-DGVSRLGSDVQSGC-VGGOAYMCDN 200
Db 1 GIGOTTRYWDCCKPSCAWDPGKP-SSPVQADKNDNPFNDGGSTRSCDAGGAYMCSQ 59
QY 201 QPWWNDLALYGAFAASLGSAGASAFCCGCVYELTFTNTAVAGKFFVQVVTNTGDDLSTNH 260
Db 60 SPWAVSDELSYGMAAVKLAGSSESQWCCACVYELTFTSGPVAGKMWIQAINTGGDLGNH 119
QY 261 FDLQMPGGVGVYFNGCOSQWNTNDGWARYGGISSISECDKLPQLQAGCKWRFGWFKN 320
Db 120 FDLAIPGGGVGIFNACTDQYGAPPNGWGDYGGIHSKEECESFPPEALKPGCNWRFDFWQ 179
QY 321 ADNPVETFKAVTCTPAEIIAKTCER 345
Db 180 ADNPVETFKAVTCTPAEIIAKTCER 204

RESULT 12

US-10-007-521-22
; Sequence 22, Application US/10007521
; Publication No. US20030054539A1
; GENERAL INFORMATION:
; APPLICANT: Schulein, Martin
; Andersen, Lene N.
; Lassen, Soren F.
; Kauppinen, Markus S.
; Lange, Lene
; Nielsen, Ruby I.
; Ihara, Michiko
; Takagi, Shinobu
; TITLE OF INVENTION: No. US20030054539A1el Endoglucanases
; NUMBER OF SEQUENCES: 109
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. US20030054539A1o No. US20030054539A1disk of No. US200300545
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/007,521
FILING DATE: 10-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/651,136
FILING DATE: 21-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4366.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 310 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-10-007-521-22

Query Match 35.3%; Score 668; DB 9; Length 310;
Best Local Similarity 57.1%; Pred. No. 4.3e-45;
Matches 117; Conservative 29; Mismatches 55; Indels 4; Gaps 3;
QY 142 SGNGRTRYWDCCKPSCAWDGKASVTKPVLTCAKDGVSRIGSDVQSGC-VGGOAYMCDN 200
Db 19 SGTGRTRYWDCCKPSCGWDKASVSQPVKTCDRNN-NPLASTARSGCDNSGVAYTCNDN 77
QY 201 QPWWNDLALYGAFAASLGSAGASAFCCGCVYELTFTNTAVAGKFFVQVVTNTGDDLSTNH 260
Db 78 QPWWNDLALYGAFAATFASGGSSEASWCCACVYELTFTSGPVAGKMWIQAINTGGDLGNH 137
QY 261 FDLQMPGGVGVYFNGCOSQWNTNDGWARYGGISSISECDKLPQLQAGCKWRFGWFKN 320
Db 138 FDLMPGGGLGIFDGTGTPQGVSYFP--GNRYGGTTSRSQCSQIPALOPCNCWRYDFWFD 195
QY 321 ADNPVETFKAVTCTPAEIIAKTCER 345
Db 196 ADNPVETFKAVTCTPAEIIAKTCER 220

RESULT 13

US-09-261-329-3
; Sequence 3, Application US/09261329
; Publication No. US20030092097A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Kim
; APPLICANT: Schulein, Martin
; APPLICANT: Christiansen, Lars
; APPLICANT: Damgaard, Bo
; APPLICANT: Von Der Osten, Claus
; TITLE OF INVENTION: Cellulase Variants
; FILE REFERENCE: 4887.204-US
; CURRENT APPLICATION NUMBER: US/09/261,329
; CURRENT FILING DATE: 1999-03-03
; EARLIER APPLICATION NUMBER: 1013/96
; EARLIER FILING DATE: 1996-09-17
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 202
; TYPE: PRT
; ORGANISM: cellulase variants
US-09-261-329-3


```
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 235 amino acids
;   TYPE: amino acid
;   STRANDEDNESS:
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   ORIGINAL SOURCE:
;   ORGANISM: Melanocarpus albomyces
;   STRAIN: ALK04237
;   FEATURE:
;   NAME/KEY: Protein
;   LOCATION: 1..235
;   OTHER INFORMATION: /label= 20K-cellulase
;
US-08-841-636A-31

Query Match      34.4%; Score 651.5; DB 1; Length 235;
Best Local Similarity 54.6%; Pred. No. 6.3e-44;
Matches 113; Conservative 27; Mismatches 64; Indels 3; Gaps 1;

Qy 139 GGFSGNGRTTRYWDCKPSCAWDGKASVTKPVLTCAKDGVSRGLSDVQSGCVGQAYMCN 198
Db 18 GALAANGQSTRYWDCKPSCGWRGKGPVNPQPVYSCDANFORIHDFAVSGCEGPAFSCA 77
Qy 199 DNQPVVYNDLAYGFAAASLGASAGAFCCGCEYELTFTNTAVAGKFFVQVNTNGDDLST 258
Db 78 DHSPWAINDLNLSYGFAATALSGQTEESWCCACVALTFTSGPVAGKTMVQVSTSTGGDLGS 137
Qy 259 NHFDLQMPGGGVYFNGCQSQWNTNTDQWARYGGISSISECDKLPLOAGCKWRFGWF 318
Db 138 NHFDLNIPIGGGVGLFDGCTPQFGGLP---GARYGGISSRQECDSFPEPLKPGCQWRFDWF 194
Qy 319 KNADNPEVTFKAVTCPAEIIAKTGCR 345
Db 195 QNADNPSFTFERVQCPPELVARTGCR 221
```

Search completed: June 18, 2003, 17:44:45
Job time : 75.2239 secs